Submission in Response to NSF CI 2030 Request for Information

DATE AND TIME: 2017-04-05 13:36:52
REFERENCE NO: 245

Author Names & Affiliations

- Lawrence Cornett - EPSCoR/IDeA Foundation

Contact Email Address (for NSF use only)

(Hidden)

Research Domain, discipline, and sub-discipline

Mammalian physiology, endocrinology

Title of Submission

Advanced Cyberinfrastructure Needs to Support Science and Engineering Research in NSF EPSCoR Jurisdictions

Abstract (maximum ~200 words).

This response to NSF 17-031 Dear Colleague Letter is being written on behalf of the 24 states as well as Guam, the Virgin Islands and Puerto Rico that are eligible for funding from the Established Program to Stimulate Competitive Research (EPSCoR) and Institutional Development Award Program (IDeA) funding. Collectively, the institutions in these jurisdictions receive approximately 10% of federal academic research dollars, yet make significant contributions to advancing the national research agenda and training the nation’s science and engineering workforce. In addition, many of these institutions address issues that are of particular importance to their home state that might not be addressed otherwise. With respect to future needs for advanced cyberinfrastructure, the institutions in the EPSCoR/IDeA-eligible jurisdictions face the same challenges and have the same needs as institutions in the rest of the country, although the challenges are often amplified in the EPSCoR/IDeA states In developing strategies to address advanced cyberinfrastructure needs to support science and engineering research in the future, the EPSCoR/IDeA Foundation encourages NSF to consider approaches that are inclusive and take advantage of the diversity that exists across the country.

Question 1 Research Challenge(s) (maximum ~1200 words): Describe current or emerging science or engineering research challenge(s), providing context in terms of recent research activities and standing questions in the field.

Cyberinfrastructure is revolutionizing many science and engineering fields, building upon capabilities developed in recent years. In fact, current capacities --or the absence of such capabilities for certain researchers and areas -- underscores the tasks ahead. Advanced cyberinfrastructure can contribute to new understanding across a wide range of areas, including extreme weather; integration of food, energy and water systems; surface and subsurface flows; ocean currents; seismic activity; fire characterization and behavior; space weather; understanding of biological systems, genomics; and materials science.

Such research not only expands knowledge but also contributes to the resolution of issues that arise daily. For example, extreme weather impacts individual lives on a daily basis, production and distribution of food supplies, transportation of goods (virtually all items in the U.S. arrive by truck), tourism and recreation, and a variety of other activities. Surface and subsurface flows affect water quality and quantity and also oil and gas exploration. Fire models inform decisions on responses to fires and safety of firefighters.

Few fields would fail to benefit from advanced cyberinfrastructure capabilities. For example, with large numbers of sensors currently deployed and multiples of these expected to come online in the upcoming years, there will be challenges to integrating individual collected data across a wide area and converting that data to useful information, such as snowmelt, streamflow related to storage and allocation. There will be opportunities to integrate a number of datasets that tell us more about drought. A similar opportunity will exist to develop a more complete understanding of the interaction of food, energy and water processes. Current wildfire models which have their limitations can be strengthened from the data collection and input stage forward as discussed at a recent National Academy of Sciences workshop entitled “A Century of Wildland Fire Research”. Better understanding of ocean currents can affect shipping, debris movement, and national security. New knowledge on the solar cycle and its impact on GPS and other technologies will be important for commercial and defense planning. Better computational hydrology models, as contemplated in Exxon-Mobil NCSA collaboration funded by NSF, can contribute to efficiency in energy exploration and energy security.

Question 2 Cyberinfrastructure Needed to Address the Research Challenge(s) (maximum ~1200 words): Describe any limitations or absence of existing cyberinfrastructure, and/or specific technical advancements in cyberinfrastructure (e.g. advanced computing, data infrastructure, software infrastructure, applications, networking, cybersecurity), that must be addressed to accomplish the identified research challenge(s).

Data-intensive science requires a continuum of technology and support stretching from data collection and input, through model development and validation, to data analytics and reproducibility. It requires an advanced computing infrastructure that is scalable and updatable, supported by significant storage of various types depending on the specific task. Furthermore significant automatic data collection currently being deployed requires the ability to deal with huge amount of data inflow that must be managed in strong metadata systems that allow metadata conversion to enable efficient computing and output control. Underlying these needs is a requirement for robust connectivity to support data inflow, and to push data to other locations that are literally worldwide. There must be both access and availability for researchers to remote datasets and scientific equipment, to colleagues who are either necessary or valuable collaborators and to the specific research community of which they are a part. Data-intensive science also requires workforce development both at the campus level and through services such as XSEDE and Jetstream. This is a major challenge, no matter how good the infrastructure, effective use of the resources requires highly trained and skilled scientists and technicians. This specific problem must be recognized and resolved across the Nation as the new economy becomes the main driver of employment.

Question 3 Other considerations (maximum ~1200 words, optional): Any other relevant aspects, such as organization, process, learning and workforce development, access, and sustainability, that need to be addressed; or any other issues that NSF should consider.

Advanced computing facilities, scientific equipment, large datasets, current models and simulations and related capacities are widely dispersed across the nation, with some researchers benefitting both from sophisticated capacity and geographic proximity to that capacity that others do not share. Individual research institutions also vary significantly in their on-site infrastructure, expertise and capability to operate and maintain cyberinfrastructure. Research institutions also vary in their connectivity, as do communities in general. Major efforts are underway to enhance connectivity through individual city and state initiatives, regional optical networks, federal networks such as ESnet, organizations such as Internet 2, and the private sector. Whatever the mechanism, it is important that major research institutions have affordable access to connectivity that will enable them to participate fully in the national research community.

Every research institution should have a plan for a 100 gigabit connection if it does not currently have such and special attention should be given to the states that participate in the NSF ESPCoR. NSF EPSCoR has a history of providing significant help in connecting institutions in EPSCoR jurisdictions to Internet2. Some EPSCoR jurisdictions have developed significant infrastructure including significant computing capacity and connectivity. Louisiana and Wyoming are excellent examples. Ongoing state investment supports the continued build-out of connectivity in the Louisiana Optical Network Initiative. a Wyoming’s Unified Network is expanding high-bandwidth connectivity to every public school in the state. These two examples demonstrate that the EPSCoR states make commitments to cyber- and computational infrastructure recognizing the imperative that the new economy means for employment and continued economic growth.

Workforce development must continue to be a high priority. As NSF has already recognized, many researchers will need assistance in accessing and utilizing major computing facilities. Campus support for researchers desiring to expand their research with computational capabilities will be needed. Additional investment in undergraduate and graduate students will be necessary. NSF should ensure that all grants that could benefit from additional computational and networking capacities allow for a contribution to the capacity. Special attention should be given to grants made to researchers in the EPSCoR states.